

# Time series modeling of Army Mission Command communication networks: an event-driven analysis

Laura R. Marusich<sup>1</sup> · Norbou Buchler<sup>1</sup>

Published online: 8 February 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** We examine the communication time series of a fully-networked Army coalition command and control organization. The network comprised two echelons of command, at the Division and Brigade levels, over a 2-week military scenario exercise involving a Mission Command staff communicating over email and phone. We used time series analysis to predict the communications record based on an external work variable of the number of important scenario events occurring across time. After taking into account structural features of the time series—decreasing communications over time, a network crash, and the transition between weeks—we examined the remaining variability in email and phone communication. We found that the exercise scenario events were not a significant predictor of the Divisional communications, which were best fit by an auto-regressive model of order 1, meaning that the best predictor of the volume of communications at a given time point was the volume of communications on the immediately preceding time point. The occurrence of scenario events, however, did predict the Brigade communication time series, which were well fit by a lag dependent variable model. These results demonstrate that Brigade communications responded to and could be predicted by battlefield events, whereas the Division communications were only predicted by their own past values. These results highlight the importance of modeling environmental work events to predict organizational communication time series and suggest that network communications are perhaps increasingly dependent upon battlefield events for lower echelons of command closer to the tactical edge.

**Keywords** Time series analysis · Communication network · Event-driven analysis · Mission Command · Dynamic networks · Social network analysis

---

✉ Laura R. Marusich  
laura.r.marusich@mail.mil; laura.r.marusich.ctr@mail.mil

<sup>1</sup> U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD 21005, USA

Advances in information and network technology continue to transform the way organizations communicate and operate, so much so that the emergence of networked forms of organization lies at the core of the economic, military, political, and social fabric of the twenty first century (Castells 1996). With the rise of networked forms of organization, recent decades have seen increased interest and research of social and organizational phenomena within real-world complex networks. Current technology enables researchers and engineers to collect, extract, and analyze large volumes of time-stamped communications (including email, text messaging, and phone correspondence) to characterize the digital behavior of individuals and the networked organization as a whole. The resulting time-series of digital communications offer rich characterizations of time-evolving networks. As in Monge and Contractor (2001), we refer to this type of network as a *communication network*, which defines the behavior of networked organizations by the flow of messages among communicators across space and time. The analysis of communication networks is understood as a multi-level and multi-scale problem that spans individual, team, and organizational levels as well as many time scales. A typical approach is to model the behavior of individual nodes within the network or the changing structural properties of the aggregate network as completely determined by the social relations and responses to other individuals within the network (e.g., Borgatti and Halgin 2011; Snijders 2005; Vázquez et al. 2006). However, there are many instances of work-directed organizations such as businesses, hospitals, police and fire departments, and military organizations that respond to and are primarily driven by external environmental events. The military command and control (C2) domain provides a clear and well-defined example of such a networked organization that requires timely communication and decision-making in response to critical collaborative work events. Within military C2 organizations, in addition to the communications network, it is common for relevant battlefield events to be digitally recorded and tracked through significant activity reports.

The objective of this work is to characterize the communication dynamics of a real-world coalition C2 organization in response to a well-defined series of significant environmental work-events. We took a time series analysis approach using data collected from an Army exercise event. Our goal was to determine how well communication volume may be predicted by the number of important environmental events occurring over time (i.e., an external work variable). Our findings indicated that within the coalition C2 network, the communication volume of a UK Brigade responded to and could be predicted by battlefield events, while the communications of the US Division were predicted only by recent communications history within that sub-organization.

## 1 Background

Many real-world complex networks: social, information, technological, biological and otherwise—exhibit common statistical, structural, and temporal properties (Newman 2003; West 2012). For example, the communication patterns of

individuals within complex social networks often exhibit “bursty” behavior (Barabási 2005), where brief periods of rapid activity are interspersed with long periods of little or no activity. Such bursty behavior leads to a heavy-tailed distribution of time intervals between an individual’s successive communications, with many very short intervals and a few extremely long intervals. One way these heavy-tailed communication patterns have been modeled, with mixed success, is with priority queuing models (Barabási 2005; Vázquez et al. 2006), in which incoming communications are assigned a priority and placed in a queue, with the individual responding to highest-priority items first. Other work has demonstrated that bursty communications can alternatively be modeled by a cascading Poisson process that incorporates circadian and weekly fluctuations in activity (Malmgren et al. 2008). Similarly, self-exciting point processes have been used to model email communication events of individuals within a corporate organization (Navaroli and Smyth 2015) and within a cohort of cadets at the West Point Military Academy (Fox et al. in press). These self-exciting point processes posit a steady background rate of email behavior with an additional self-exciting function which leads to clustering of email activity.

These approaches have primarily focused on the level of the individual within a network, rather than on the dynamic behavior of the organization as a whole. Furthermore, several of these approaches consider an individual’s response to a steady stream of incoming communications, but they do not take into account the effects of external work events that we posit can drive an organization and influence the rates of both incoming and outgoing communications. Our approach examines an evolving time series of communications of a hierarchical networked organization, focusing in particular on the effects of external environmental events.

Understanding how an organization functions requires examining the environmental context of the organization (see Ancona and Caldwell 1992). This argument is not new. In a seminal book, *The Sciences of the Artificial*, Nobel laureate Simon (1996) offers the parable of an “ant wandering on a pebbled beach” to suggest that the apparent complexity of human behavior might be due to environmental factors:

An ant, viewed as a behaving system, is quite simple. The apparent complexity of its behavior over time is largely a reflection of the complexity of the environment in which it finds itself. I should like to explore this hypothesis with the word ‘man’ substituted for ‘ant’.

Theorists in the ‘grounded cognition’ movement (Clark 1997; see Barsalou 2010 for a review) have recently adopted this argument by noting that much of the dynamics of human behavior stem from reciprocal causation links to the environment as well as to other humans (e.g., economic systems), and thus the environmental context must be fully considered. Here we apply this perspective by attempting to understand the dynamics of communication within a networked organization not just as a function of internal structure and states, but also in relation to the environmental context of external driving events. We use time series analysis techniques to examine the record of email and phone communications in a military C2 network to discover the extent to which we can predict the inter- and intra-

organizational communications time series given a second time series of significant work events.

## 2 Time series data and regression analysis

The general purpose of regression analysis, whether used in time series or cross-sectional data, is to determine the existence and form of the relationships between variables. Regression analysis typically shows how some dependent variable (DV) is affected by changes in one or more independent variables (IVs).

In our examination of networked communications, we use the timing of work events as an IV, and the communications time series as the DV. A key consideration is that the network communications reflect both the work process itself and the dissemination of both intermediate and final work products as they occur across the communication network over time.

There is some evidence to suggest that organizations increase communications in response to critical events. In a detailed analysis of the email corpus of the Enron Corporation, (Diesner et al. 2005; see also Murshed et al. 2007) found that during the crisis period of financial insolvency, the volume of communications intensified among employees, becoming more diverse with respect to established contacts and formal roles. The Enron crisis is instructive as a communications network both before and during a critical period of time. Other researchers have also found dramatic changes in email usage following major (typically singular) environmental events in organizational networks, such as after a corporate merger (Danowski and Edison-Swift 1985) or after downsizing (Shah 2000). Remaining unexamined is whether communications can be understood in response to more regular and recurrent work events, as opposed to singular events. In our study, the work events are many and varied in type and urgency of response, and they are formally established as part of a scenario in a military training event exercise. We used social network analysis and time series analysis techniques to determine the characteristics of communications in the hierarchical network (i.e., a Division and subordinate Brigade staff) during the scenario-based training event. We seek to determine whether the communication time series can be explained and/or predicted on the basis of a time-line of work events.

## 3 Method

The communications and scenario-event data described in this paper were collected at a 2-week US Army simulation-based training event. The Mission Command Battle Laboratory at Fort Leavenworth, Kansas conducted a joint experiment/exercise—with aspects of both an experiment and a training exercise—assessing the operations of the Mission Command staff composed of a US Division headquarters ( $n = 51$ ) and subordinate UK Brigade headquarters ( $n = 28$ ). Two additional US Brigade Combat Teams participated in the exercise; however, the scenario focused primarily on the Division and the UK Brigade, and these US Brigades were only

partially staffed. As a result, we limit our analysis to the US Division and UK Brigade sub-organizations. The exercise involved the simulation of a continuous 48 h of scenario events; however the exercise was “active” for 6-h periods on Monday–Thursday and “paused” between these periods. As a result, participants completed the 48 h of scenario time over two “real-world” weeks. The network architecture and digitized nature of the event allowed examination of staff communications in a distributed, network-enabled coalition environment. The participants were active duty (and in some limited cases, retired) soldiers and officers, organic to their military unit. All participants either currently served in the role they took during the scenario, or had extensive past experience in that assigned role. Below we describe the defining characteristics of this military ad-hoc organization, and of the tasks they were required to complete.

### 3.1 Defining characteristics of the organization

- (1) *Real organization* several groups participated in the exercise, including a representative C2 headquarters of a US Division, a fully-staffed UK Brigade, and two partially-staffed US Brigade Combat Teams. The participating organizations were existing units, whose staff execute differentiated, well-specified, and interdependent roles. The units operated in a distributed-fashion (US units in Fort Leavenworth, KS and UK staff in London) over a communications network using specialized military C2 hardware and software.
- (2) *Convened to accomplish a particular training mission* the military organization was staffed and convened specifically to execute and accomplish a particular 2-week long training mission.
- (3) *Common mission* members are interdependent in their work and decision-making about how to proceed with the mission. The organization functions as a *purposive social system*, where members are readily identifiable to each other by role and work interdependently to accomplish one or more collective objectives (Hackman 2008; Hackman and Katz 2010). The responsibility for performing the various tasks and sub-tasks necessary for mission success is divided and assigned among the staff.

### 3.2 Defining characteristics of the tasks

The training scenario in a military exercise generates many overlapping series of event-driven tasks, the resolution of which requires a high degree of coordination among the participating C2 staff. The training scenario involved a coalition environment with a UK Brigade operating under the command of a US Division, focused on coalition interoperability. Two US Brigades also operated under the command of the US Division, but neither was fully staffed nor fully engaged by scenario events. The scenario focus was on the networked organizational communications within and between the US Division and UK Brigade. Researchers

have long pointed out that the nature of a task has a great influence on the steps and processes a group uses to perform the work (e.g., McGrath and Kravitz 1982; Roby and Lanzetta 1958). The tasks of groups in the military domain considered here have four distinguishing features:

- (1) *Specific presenting problems* the military C2 staff must address specific problems that occur in the unit's area of operations. The military staff organization must monitor key events and successfully plan and coordinate an effective response, given limited resources. The presenting problems may be kinetic events, such as an improvised explosive device, or civil–military in nature, such as responding to a civil demonstration. At other times, the presenting problem may be a time-sensitive intelligence report of enemy activity that needs to be analyzed and corroborated. At any given time, the organization must coordinate a response to many such presenting problems.
- (2) *Adherence to specific tactics, techniques, and procedures* the groups adhere to formalized military work routines and processes that are known in advance, delegating specific work responsibilities for the various sub-groups and individual members.
- (3) *Addressed immediately* the group operates in an urgent, time-sensitive work environment and must immediately coordinate responses to work events that may have adverse cascading effects if not addressed in a timely manner.
- (4) *Results in collaborative work products that need to be coordinated and disseminated* the group is expected to construct specific, detailed material products that will exist independently of the group process or the individual members themselves. For instance, the Commander and his command elements require regular reports from the staff in order to achieve situational awareness of the battlefield environment. The work process itself and the dissemination of both intermediate and final work products occur across the communication network over time.

### 3.3 Data collected

#### 3.3.1 Communications

Telephone (VoIP) and email were the two primary methods of communication between staff members during the exercise. For each phone call made and email message sent in our dataset, three pieces of information were automatically electronically logged: the sender, the receiver, and the time of the communication's initiation.

#### 3.3.2 Scenario events

We also collected information about the various scenario events, including: the time at which each scenario event occurred, and the group that was affected by each event. The specific timing of the significant events was known to the researchers

ahead of time, but not to the military units participating in the training exercise. Knowing the scenario event timeline is advantageous to researchers interested in examining the correspondence between specific work events (i.e., presenting problems) and the time series of communications.

### 3.4 Data analysis: social network analysis

The first step in analyzing this communications dataset was to construct and visualize it as a directed network representing the total communications that occurred across the entire exercise. Each node in the network represents an individual, and each link represents at least one communication event between two individuals during the exercise, with thicker links indicating a greater number of communication events.

While each individual in this network was free to communicate with any other, an important feature of the network is that it is composed of two groups differing in both echelon (the Brigade is a subordinate element of the Division) and nationality (US vs. UK). Thus we are interested in both the overall number of links and whether those links connect individuals in the same or different groups. Figure 1 shows the network broken out to depict the internal and external links. There were 51 individuals in US Division roles and 28 in UK Brigade roles, comprising the total 79 nodes in the network.

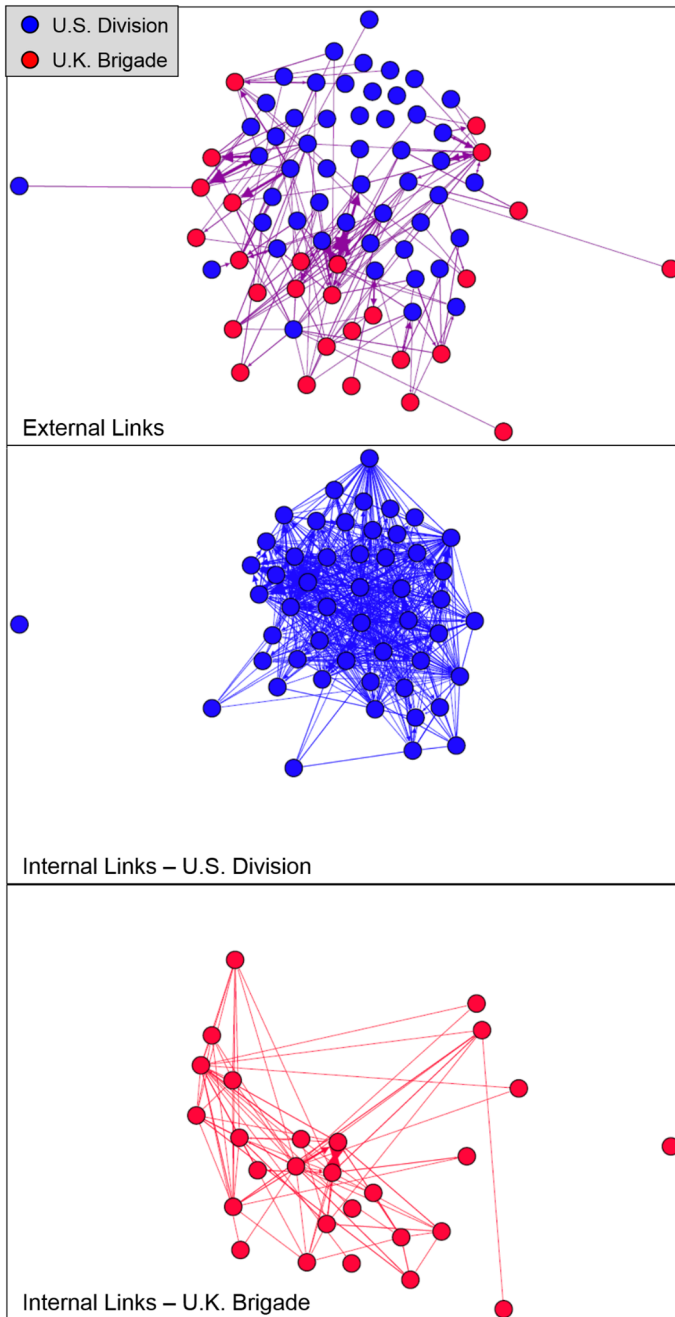
This visualization shows more internal than external communications and denser internal communications in the US Division group than in the UK Brigade. As in most real-world networks, the distribution of communication links among the nodes is highly skewed, with a small subset of nodes having many connections, while most nodes have a small number of connections.

Table 1 shows the value of several structural network properties, averaged across the US Division nodes and the UK Brigade nodes separately. On average, US Division members had more links to other nodes (higher degree), shorter path lengths to other nodes (lower closeness centrality), more instances of acting as bridge between two other nodes (higher betweenness centrality), and more dense clustering (higher clustering coefficients). These results indicate that the members of the US Division were, on average, more central to the overall network and more densely connected than members of the UK Brigade.

### 3.5 Data analysis: time series analysis

The above analyses provide an interesting depiction of the overall distribution of communications within and between groups in the US Division and UK Brigade; however, they do not address time. Analyzing the communication events and network statistics as time series allows us to explain the changing dynamics of the collaboration and possibly forecast future communication patterns.

The time series analysis first required aggregating the data into discrete temporal intervals. The appropriate interval size over which to integrate depends on the overall time span and resolution of the data in question. We chose 1-h intervals to allow for both a sufficient amount of variability between intervals and a large



**Fig. 1** Directed network of total communications across the entire exercise, separated into (1) external links between the US Division and UK Brigade, (2) internal links within the US Division, and (3) internal links within the UK Brigade. Edge width represents number of communication events (emails or telephone calls) between staff members. Arrows point from communication senders to receivers



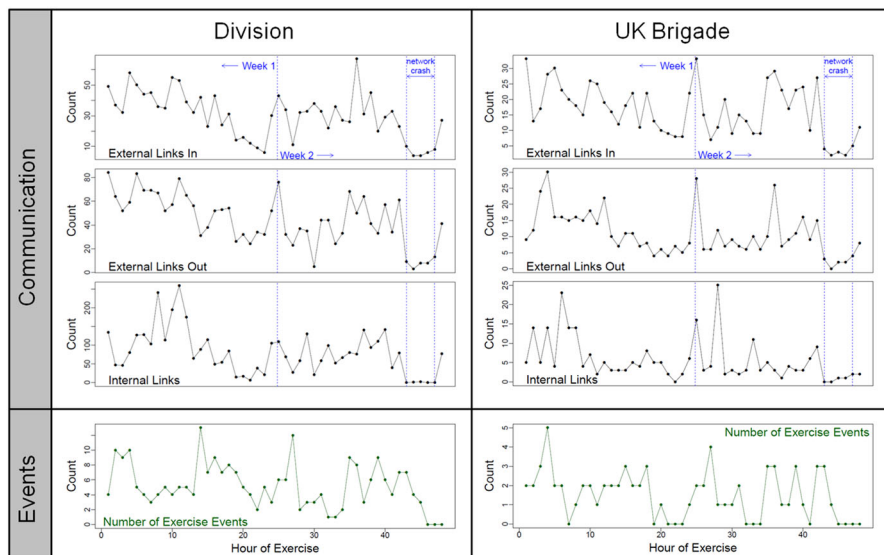
**Table 1** Average structural properties of US Division and UK Brigade nodes

Properties	US Division (mean)	UK Brigade (mean)
Degree	36.31	15.43
In-degree	17.80	8.36
Out-degree	18.51	7.07
Closeness centrality	1.76	1.99
Betweenness centrality	86.03	47.19
Clustering coefficient	0.55	0.30

enough number of time points ( $t = 48$ ) while avoiding too many empty intervals. Thus, we aggregated three variables into intervals of 1 h for each of the two groups:

- (1) *External links in* the number of links directed from an individual outside of the group to an individual within the group.
- (2) *External links out* the number of links directed from an individual within the group to an individual outside of the group.
- (3) *Internal links* the number of links directed from an individual within the group to another individual within the group.

These three variables are plotted against time for the two groups in Fig. 2. Also noted on this figure is the time point that marks the start of the second week of the exercise, as well as the five time points during which there was a planned



**Fig. 2** Time series of communications and scenario events for the US Division and the UK Brigade over 48 h of the military exercise

degradation of the computer network. One effect of this degradation was to prevent all electronic communication; however, phone communications were still possible. Below these three communication variables, we plot the time series (aggregated by hour intervals) of the number of scenario events that affected both groups.

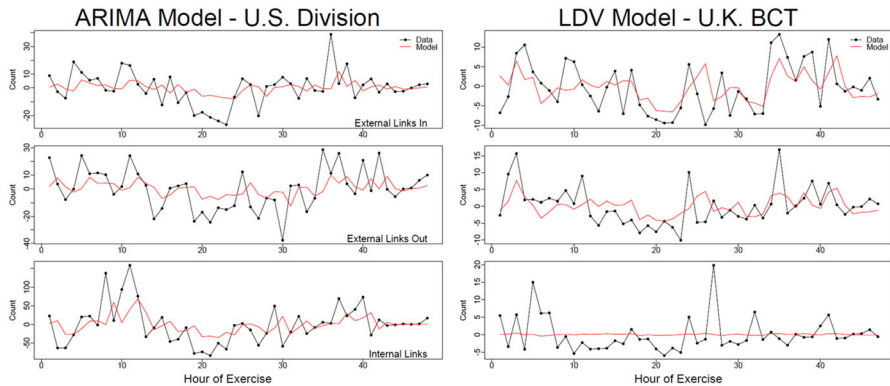
We are interested in predicting the fluctuations in volume of communication links using the number and timing of the exercise events. However, time series data require additional considerations than are typically necessary for a straightforward regression analysis with cross-sectional data. Random sampling from the larger population is assumed in typical cross-sectional data, such that each case is independent of the others. In time series data, sequential observations of a single case are collected. Instead of each observation representing a random draw from a larger population, the sequence of observations represents one possible realization of a stochastic data generating process. The corresponding assumption to random sampling when using time series data is stationarity: the expected value of the variable of interest must be constant over time.

Assuming the variables are stationary, the sequential observations that are collected in time series data are typically not independent, but are often serially correlated. One of the assumptions required by ordinary least squares regression is that the residuals be independent of each other (i.e., the magnitude of one error must not be predictable from the others). If there is serial correlation in the DV, and it is not accounted for by the explanatory variable(s), this assumption will be violated. As a result, past values of the DV are often included as predictors in time series analysis models.

### 3.5.1 Stationarity: trending and structural breaks

As described above, one of the key conditions of time series analysis is that the variables be stationary. Figure 2 suggests that our raw communications time series may not meet this assumption. Many of the sequences appear to display a downward trend with time. In addition, two features of the exercise seem likely to exert an influence on the values of the DVs. The first is the transition between Weeks 1 and 2 of the exercise. The second is a nearly 5-h time period during which a planned computer network degradation occurred. Both of these events might be expected to shift the mean of the DV to a different value at the times they occurred, which is a violation of stationarity known as a *structural break*. Trending and structural breaks can be included in a model along with exogenous variables; however, they can artificially improve fit statistics, and so it is often more appropriate to remove their effects before fitting a model. This is done by simply regressing the DV on these variables, and using the residuals from this regression as the new DV of interest. This new DV will have a mean of zero, and if there were no other factors causing nonstationarity, will be stationary.

Accordingly, we regressed our three communication variables on a linear time variable and two dummy variables, one representing the transition between weeks and one representing the network degradation. The residuals from this regression (Fig. 3) represent the variability in the communications variables that are not explained by a linear trend with time or the two structural variables.



**Fig. 3** Detrended data (filled black circles) and model predictions (red lines). (Color figure online)

We also detrended the events time series, as they also showed evidence of an overall decrease with time. Visually, the detrended data do not show obvious evidence of nonstationarity. In addition, unit root tests performed on the detrended data do not indicate stationarity violations. The models described below were all fit using the detrended communications and event time series.

## 4 Results

### 4.1 Time series models

There are many varieties of models to fit time series data. We were guided first by theory in choosing among these options (i.e., what is expected to affect the volume of communications in this exercise) and second by parsimony. Our goal in fitting time series models to the data was to determine whether the volume of communications could be predicted based upon previous communications and/or the dynamic volume of scenario events encountered, and whether communications are predicted differently in the US Division than in the UK Brigade. We started by examining models that included up to two lags of the DV and the IV, as it is plausible that events and communications from 2 h previous could be predictive of communications in the current hour. Terms that were not significant for any of the DVs were removed, and the Akaike information criterion was assessed as a measure of model fit. In the end, the simplest model that provided the best fit for most of the DVs in the UK Brigade was a lag DV (LDV) model with one lag of the DV (communication) and only the contemporary values of the IV (scenario events) included as predictors. In the US Division, the best model was an autoregressive model of order 1 [AR(1)]. These two types of models are described in detail below.

#### 4.1.1 Lag dependent variable model

The LDV model differs from simple linear regression in that it includes one or more lags of the DV in its set of predictors, meaning that previous values of the DV are used to predict the current value. With one IV and a single lag of the DV, it takes the following form:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta x_t + \varepsilon_t,$$

where  $y_t$  is the time series to be predicted,  $\alpha_0$  is a constant,  $\alpha_1$  is a parameter that determines the weight of the previous observation of  $y$ ,  $x_t$  is the explanatory variable,  $\beta$  is the weighting parameter on  $x_t$ , and  $\varepsilon_t$  is an uncorrelated noise process. The inclusion of the lagged terms makes interpretation of the coefficients on the IVs somewhat more complicated than simple linear regression where the interpretation of the slope parameter is straightforward; it is the immediate change in the value of  $y$  predicted by a one-unit change in  $x$ .

An analogous understanding of the relationship between variables in the LDV can be obtained by calculating the short and long run effects of the IV upon the DV. The short run effect is the immediate change in the DV caused by a temporary one-unit increase in the IV. If  $x$  were to increase by 1 for a single time point and then return to its prior value in the next, the short run effect is the corresponding immediate change in  $y$ . It is equal to  $\beta$ , the coefficient on the IV. The long run effect is the asymptotic or eventual change in the DV caused by a permanent one-unit increase in the IV. If  $x$  were to increase by 1 and stay at that new value for all future time points, the long run effect is the corresponding eventual change in  $y$ . It is calculated by dividing  $\beta$  by one minus  $\alpha_1$ , the coefficient on the lagged DV. Analogous to the intercept term in a simple regression, the equilibrium of the process is calculated by dividing  $\alpha_0$  by one minus  $\alpha_1$ .

Figure 4 depicts the relationship between the two coefficients and the short and long run effects in the LDV. The top panel is a plot of the IV over time, with one transient one-unit increase and one permanent one-unit increase. The middle panel shows the corresponding changes in the DV over time, using a constant value for  $\beta$  and three different values of  $\alpha_1$ . In the bottom panel,  $\alpha_1$  is held constant, and  $\beta$  is allowed to vary. This figure shows that  $\beta$  determines the amplitude of both the short and long run effects, while  $\alpha_1$  determines the time course of the effects, as well as the amplitude of the long run effect.

#### 4.1.2 Autoregressive moving average model

We did not find significant contributions of the scenario events to the communication variables, whether we used contemporary or lagged values in the US Division data. There were, however, significant autocorrelations in the DVs. As a result, we found that fitting an AR moving average (ARMA) model to the data provided a good explanation of the dynamics. An ARMA(1, 1) model takes the general form:

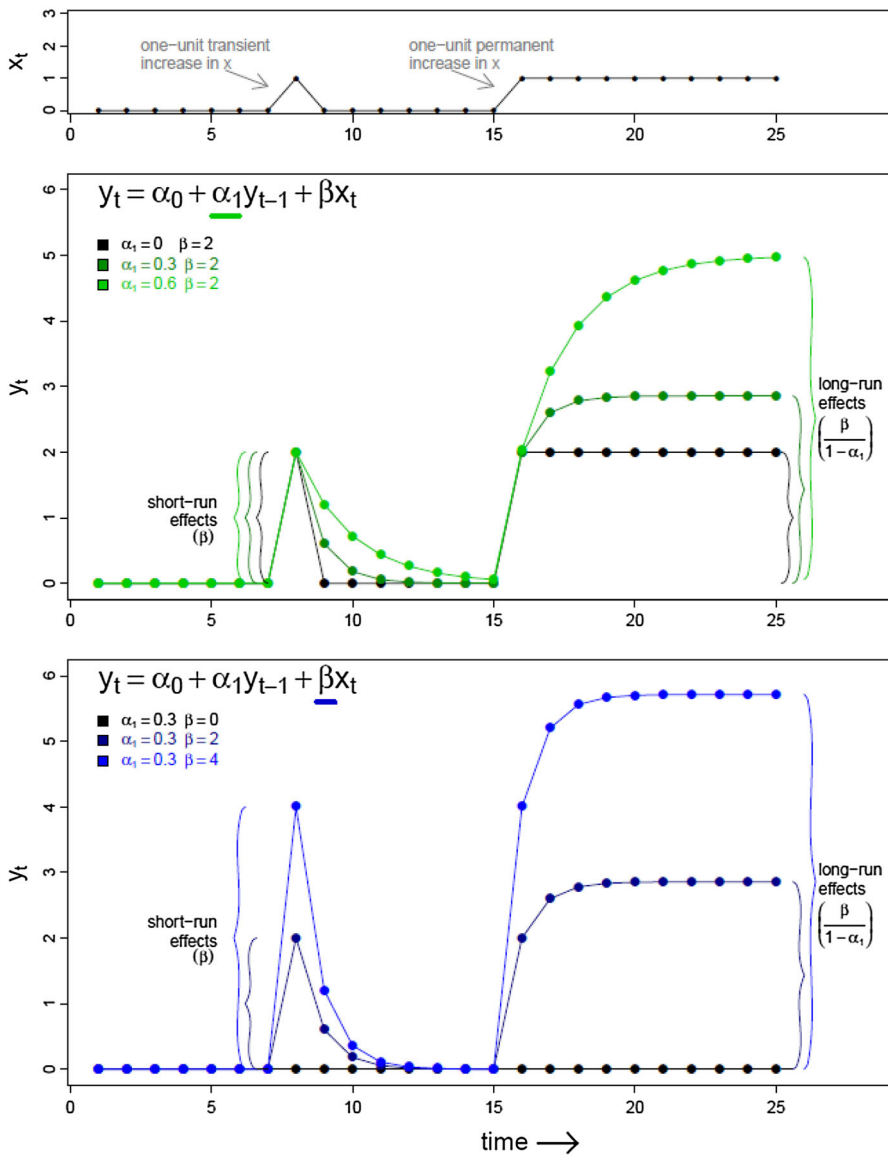


Fig. 4 Short- and long-run effects in the lag dependent variable model

$$y_t = \alpha_0 + \sum \alpha_i y_{t-i} + \sum \phi_i \varepsilon_{t-i} + \varepsilon_t,$$

where  $y_t$  is the time series to be predicted,  $\alpha_0$  is a constant,  $\alpha_i$  are parameters that determine the weighting of past observations of  $y$ ,  $\varepsilon_t$  is an uncorrelated noise process, and  $\phi_i$  are parameters that determine the weighting of past values of  $\varepsilon_t$ .

No external regressors are included, and so this model only predicts future values based upon internal dynamics.

We used the Box–Jenkins method for selecting the appropriate ARMA model, finding that the US Division data is best explained by an AR(1) model, meaning that it includes one lag of the DV and no moving average term. This takes the form:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t.$$

In this model, the DV is predicted only from its own previous behavior, and because the model is AR(1), only the immediately previous value of  $y$  is used to predict the current value. In contrast to the LDV model, the understanding of short and long run effects of the IV upon the DV does not apply in this model because there are no external IVs included as predictors. The process equilibrium is calculated by dividing  $\alpha_0$  by one minus  $\alpha_1$ .

## 4.2 Model predictions

The best model of the UK Brigade communications is the LDV, which includes both the lagged communications and the event time series as predictors. The best model of the US Division communications is an AR(1) model, which only includes the lagged communications as a predictor. The de-trended communication time series and the predictions of the best model for each group are plotted in Fig. 3.

The parameters for the best-fitting model for each group, as well as the model's root mean squared error (RMSE), appear in bold in Table 2. For comparison's sake, we also show the results of fitting an LDV model to the US Division data and an AR(1) model to the UK Brigade data. The US Division LDV model results show that the coefficient on the independent events variable ( $\beta$ ) is never significant. Further, including the added events predictor does not reduce the RMSE, supporting the AR(1) as the best model for the US Division data. Conversely, most of the UK

**Table 2** The parameters for the best-fitting AR(1) model and LDV model are shown for both the US Division and the UK Brigade, along with the root mean square error for each model

Communication variables	AR(1)			LDV			
	$\alpha_0$	$\alpha_1$	RMSE	$\alpha_0$	$\alpha_1$	$\beta$	RMSE
US Division							
External links in	<b>0.11</b>	<b>0.31*</b>	<b>11.32</b>	−0.17	0.31*	0.08	11.37
External links out	<b>0.34</b>	<b>0.34*</b>	<b>13.98</b>	−0.39	0.33*	−0.31	13.75
Internal links	<b>0.60</b>	<b>0.43***</b>	<b>45.82</b>	−0.16	0.43**	−2.91	45.46
UK Brigade							
External links in	0.08	0.29*	6.26	<b>−0.30</b>	<b>0.25*</b>	<b>2.53***</b>	<b>5.27</b>
External links out	−0.03	0.24	5.43	<b>0.12</b>	<b>0.18</b>	<b>2.00**</b>	<b>4.90</b>
Internal links	0.00	−0.02	4.91	<b>0.08</b>	<b>−0.01</b>	<b>0.14</b>	<b>4.93</b>

The best model for each group is shown in bold

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

**Table 3** Short and long run effects in LDV model of UK Brigade communications

Communication variables	Short run effects	Long run effects
External links in	2.53***	10.71**
External links out	2.00**	8.85**
Internal links	0.14	0.05

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Brigade communications data are significantly predicted by the events variable, and the RMSE is reduced in the LDV model over the AR(1) model, supporting the LDV model in this case.

The intercept ( $\alpha_0$ ) values in both models are not particularly meaningful by themselves, considering they reflect de-trended, rather than raw data. In the AR(1) model, the  $\alpha_1$  values reflect the amount of influence the value of the previous time step has on the current one. These values are significant and in the range of approximately 0.3–0.4 for all three of the communication variables. The predicted volume of communication at any given point in time resembles a memory growth/decay function that goes up or down depending on the immediately prior communications volume with a step-size of roughly 30–40 % of that prior value.

The  $\alpha_1$  and  $\beta$  parameters of the LDV model are easier to interpret in terms of long and short run effects. The computed values for these effects are reported in Table 3. For external links in and external links out, the short run and long run effects are significant. The interpretation of these effects is the following: a temporary increase of one additional scenario event is expected to lead to 2.53 additional links directed into the UK Brigade at that time point, and 2.00 additional links directed out of the UK Brigade. A permanent increase of one additional scenario event is expected to lead to 10.71 additional links in and 8.85 additional links out.

In the AR(1) model of the US Division communications data, the internal links show the largest AR coefficient, meaning that the previous value of the volume of links exert a larger influence on the current value. Conversely, in the LDV model of the UK Brigade data, the internal links variable is the only one without a significant short or long run effect of scenario-events. The previous value of the number of internal links also did not significantly predict the current value in the UK Brigade. This is clearly seen in Fig. 3, where the model predictions capture only the mean and not the fluctuations over time.

## 5 Discussion

The ability to model and predict the communications of the US Division and UK Brigade with any degree of success is in itself interesting, apart from the actual form and interpretation of the models. The data originated from a complex network of many individuals with distinct roles and the potential to be affected by many different varying factors, both recorded (e.g., scenario events) and unrecorded.

There is no guarantee that the communications on one time point would be a significant predictor of the next time point, but in both groups (US Division and UK Brigade), the best models include the number of communication links on the previous time point as a predictor. The number of communication links at each time point was not independent of the others, but were serially correlated (i.e., high numbers of communication links tended to be followed by high numbers of communication links, and vice versa). Within the US Division, communication volume was best fit by an AR(1), meaning that the best predictor of the volume of communications at a given time point was the volume of communications on the immediately preceding time point.

In addition to serial correlation, the external UK Brigade communications time series (links-in and links-out) were also well predicted by the number of scenario events occurring within the same time period. These data were best fit by a LDV model. The finding that the UK Brigade communications responded to and could be predicted by battlefield events is also nontrivial, as the scenario events varied greatly in terms of severity and the nature of the response required, and it is not obvious that a simple measure like the number of events could predict the communications. One interesting caveat is that the internal links within the UK Brigade were not well-predicted by the scenario events. It may be the case that communications within this organization are more reflective of prescribed tactics, techniques, and procedures than they are of external circumstances, a hypothesis that may be explored in future work.

Using environmental events to predict network dynamics is not a standard approach in the social network analysis literature. Typical goals for time series analyses of network behavior are to identify and visualize significant changes in a network. Models of change often focus on the node level, where changing states of individual nodes leads to changes in the structure of the network over time (e.g., Carley 2003; Snijders 2005). Within the domain of mission control, however, researchers have examined communication volume in response to work variables such as technology utilization in distributed team operations (Caldwell 2005). Analogous to our Mission Command context, team performance in the execution of space missions is largely based on managing, and coordinating a responses to external events rather than natural cycles of group communication processes (see McGrath and Hollingshead 1994). In this work, Caldwell found that communication volume and channel utilization by the crew varied significantly across the phases of the space mission-ascent, orbit, and entry (highest during ascent). This conforms to our results that communications vary by the task demands of the external mission context. Caldwell defines the mission context by the phase of the mission, whereas we specifically define the mission context as a tick timeline of stimulating scenario events and focus on understanding communication networks in the aggregate. Our results suggest that this approach could be a valuable tool for understanding network dynamics.

Prior to the analysis, we de-trended the time series data by using a linear time variable and by removing the effects of two structural breaks (the transition between scenario weeks and the network degradation period). We performed this preprocessing step in order to meet assumptions of stationarity—a variable that trends is



not stationary, as its mean is changing with time. Similarly, stationarity is not met when the mean of a variable shifts to a new level at a particular time (e.g., the communication data during the network degradation). From this perspective, the three variables that we used to de-trend the data could be considered merely nuisance variables.

However, the significant effects of these three variables upon the communications time series provide some interesting information about our data. For example, the communication variables showed significant decreases with time. This suggests the possibility of a type of training effect, where the participants initially established communication links with many different individuals, but pruned these connections over time as they learned which links were most efficient for accomplishing their goals. As a result, the number of internal and external links decreased over time. Secondly, there was a significant increase in communications at the transition between weeks, indicating a type of reset, where the steady decrease in links over time was temporarily offset after a break and change in the mission of the exercise. Finally, there is a significant decrease in communications during the period when the computer network was degraded. This is an expected effect and provides a good check that our time series analysis methods are sensitive to the dynamics of this exercise.

One of the primary findings from our analyses is that scenario events can predict the external communication time series from the UK Brigade but not the communication time series from US Division headquarters. This finding suggests that as the focus of analysis moves further down the hierarchy and closer to the tactical edge, the communication network exhibits a stronger response to external events. This is supported by Anthony's (1999) work noting scaling relationships between echelon level and the reaction and deliberate planning cycles, with cycle times increasing proportionately with echelon level. The C2 concept of "agility" is closely aligned with the notion of responsiveness to external events. C2 agility is "the ability to successfully effect, cope with, and/or exploit changes in circumstances", and can be broken down into two categories, passive agility and active agility (Alberts 2011). Passive agility is the ability of an organization to successfully operate under a specific range of circumstances without taking intervening action. Active agility occurs when an organization takes anticipatory or responsive action to successfully operate within changing circumstances. The implication is that an organization that exhibits active agility will be able to successfully operate across a more extensive range of circumstances than one with only passive agility. Our findings provide some evidence that the UK Brigade demonstrated active agility, as this group's communication dynamics varied in response to the changing circumstances of the scenario events. Conversely, the US Division's behavior aligned more closely with passive agility, as this group's communications did not show responsiveness to scenario events. These results align with the traditional understanding of the military hierarchy, with lower echelons that are closer to the tactical edge interacting directly with the operating environment and its changing state of affairs. However, as the military increases its understanding of C2 agility, there is an increasing push to achieve active agility throughout the hierarchy (e.g., Casey 2009; Defense Advisory Committee 2012; NATO SAS-85

2013). Our approach has the potential to provide a method for assessing active agility in a network and pinpointing sub-organizations where improvement could occur.

In the overall network spanning the entire exercise period, we found several structural differences between nodes from the two echelons. On average, nodes from the US Division had more connections to other nodes in the network, were more central in the network, and tended to form more tight clusters with other nodes than individuals in the UK Brigade. These structural differences are consistent with the differences in temporal dynamics between the two groups. The implication is that the higher-echelon US Division staff are fairly constrained by prescribed workflows (e.g., regular ‘battle rhythm’) such as reporting requirements and briefing schedules that ensure a fairly high rate of regular communication among staff members, independent of external events. The structural evidence suggests that the more agile UK Brigade, on the other hand, had more limited communications that were perhaps focused more directly on responding to the scenario events in the exercise. The nature of the relationship between structural properties and temporal dynamics may be further explored in future work.

To make interpretations and predictions about humans and the organizational and environmental context in which they operate, network approaches need to capture and incorporate both performance and environmental models. In our work we combined techniques from social network analysis and time series analysis to predict the communication dynamics of a networked coalition C2 organization. Our approach was novel in that it incorporated external events and assessed where in the organization communications were driven by these events. This type of approach may be useful for ongoing and future work exploring the relationships between networked communication dynamics and objective measures of organization performance (e.g., Buchler et al. 2013, 2014; Orvis et al. 2013; Saavedra et al. 2011). A more complete understanding of the driving factors of communications within C2 networks can greatly improve such models relating communications to performance and effectiveness. There is a clear advantage to understanding how the pattern of networked communication influences the effectiveness of the organizational response to environmental events, that can lead the way for opportunities for technological and training intervention to improve effectiveness.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Alberts DS (2011) The agility advantage: a survival guide for complex enterprises and endeavors. DoD Command and Control Research Program
- Ancona DG, Caldwell DF (1992) Bridging the boundary: external activity and performance in organizational teams. *Adm Sci Q* 37:634–665

- Anthony RW (1999) Relating large and small in C2 and operations. In: Command and control research and technology symposium (CCRTS), Newport, RI
- Barabási AL (2005) The origin of bursts and heavy tails in human dynamics. *Nature* 435:207–211
- Barsalou LW (2010) Grounded cognition: past, present, and future. *Top Cogn Sci* 2:716–724
- Borgatti SP, Halgin DS (2011) On network theory. *Organ Sci* 22:1168–1181
- Buchler N, Marusich LR, Bakdash JZ, Sokoloff S, Hamm R (2013) The Warfighter Associate: objective and automated metrics for Mission Command. In: 18th ICCRTS
- Buchler N, Marusich LR, Sokoloff S (2014) The Warfighter Associate: decision-support software agent for the management of intelligence, surveillance, and reconnaissance (ISR) assets. In: Proc. SPIE 9079, ground/air multisensor interoperability, integration, and networking for persistent ISR V (pp 907902–907902)
- Caldwell B (2005) Multi-team dynamics and distributed expertise in mission operations. *Aviat Space Environ Med* 76(Supplement 1):B145–B153
- Carley KM (2003) Dynamic network analysis. In: Dynamic social network modeling and analysis: workshop summary and papers (pp 133–145). Committee on Human Factors, National Research Council
- Casey GW Jr (2009) The Army of the 21st century. *Army Mag* 59:25–40
- Castells M (1996) The rise of the network society: the information age: economy, society and culture, vol 1. Blackwell, Cambridge
- Clark A (1997) Being there: putting brain, body and world together again. MIT Press, Cambridge
- Danowski JA, Edison-Swift P (1985) Crisis effects on intraorganizational computer-based communication. *Commun Res* 12:251–270
- Defense Advisory Committee (2012) A new US defense strategy for a new era: military superiority, agility and efficiency. Stimson Center, Washington D.C.
- Diesner J, Frantz TL, Carley KM (2005) Communication networks from the Enron email corpus “It’s always about the people. Enron is no different”. *Comput Math Organ Theory* 11:201–228
- Fox EW, Short MB, Schoenberg FP, Coronges KD, Bertozzi AL (in press) Modeling email networks and inferring leadership using self-exciting point processes. *JASA*
- Hackman JR (2008) What is this thing called leadership? In: Harvard Business School centennial colloquium on “Leadership: Advancing an Intellectual Discipline,” pp 10–11
- Hackman JR, Katz N (2010) Group behavior and performance. In: Fiske ST, Gilbert DT, Lindzey G (eds) *Handbook of social psychology*, 5th edn. Wiley, New York, pp 1208–1251
- Malmgren RD, Stouffer DB, Motter AE, Amaral LAN (2008) A Poissonian explanation for heavy tails in e-mail communication. *PNAS* 105:18153–18158
- McGrath JE, Hollingshead AB (1994) Groups interacting with technology. Sage, Newbury Park
- McGrath JE, Kravitz DA (1982) Group research. *Annu Rev Psychol* 33:195–230
- Monge PR, Contractor NS (2001) Emergence of communication networks. In: Jablin FM, Putnam LL (eds) *The new handbook of organizational communication*. Sage, Thousand Oaks, pp 440–502
- Murshed T, Davis J, Hossain L (2007) Social network analysis and organizational disintegration: the case of Enron Corporation. In: ICIS2007
- NATO SAS-085 (2013) C2 agility report. STO-TR-SAS-085
- Navaroli N, Smyth P (2015) Modeling response time in digital human communication. In: Ninth international AAAI conference on web and social media
- Newman MEJ (2003) The structure and function of complex networks. *SIAM Rev* 45:167–256
- Orvis KL, DeCostanza AH, Duchon A (2013) Developing systems-based performance measures: a rational approach. In: The interservice/industry training, simulation and education conference (I/ITSEC, no. 1)
- Roby TB, Lanzetta JT (1958) Considerations in the analysis of group tasks. *Psychol Bull* 55:88
- Saavedra S, Hagerty K, Uzzi B (2011) Synchronicity, instant messaging, and performance among financial traders. *PNAS* 108:5296–5301
- Shah PP (2000) Network destruction: the structural implications of downsizing. *Acad Manag J* 43:101–112
- Simon HA (1996) The sciences of the artificial. MIT Press, Cambridge
- Snijders TAB (2005) Models for longitudinal network data. In: Carrington P, Scott J, Wasserman S (eds) *Models and methods in social network analysis*. Cambridge University Press, New York, pp 215–247

- Vázquez A, Oliveira JG, Dezső Z, Goh KI, Kondor I, Barabási AL (2006) Modeling bursts and heavy tails in human dynamics. *Phys Rev E* 73:036127
- West BJ (2012) *Complex worlds: uncertain, unequal and unfair*. Black Rose Writing, Castroville

**Laura R. Marusich** is a Postdoctoral Research Fellow in the Cognitive Sciences Branch of the US Army Research Laboratory (ARL) Human Research and Engineering Directorate (HRED). Her current research is focused on the temporal dynamics of effective collaboration. She received her PhD in Psychology from the University of Texas at Austin in 2011.

**Norbou Buchler** is a Scientist in the Cognitive Sciences Branch of the US Army Research Laboratory (ARL) Human Research and Engineering Directorate (HRED). He received his Doctoral Degree from Syracuse University in Experimental Psychology after which he did two post-docs, first at Carnegie Mellon University using computational modeling and behavior approaches to understanding cognition and the human memory system, his second post-doc at Duke University in cognitive neuroscience used fMRI approaches to elucidate the neural correlates of memory retrieval. In 2009, he joined ARL and is currently engaged in basic and applied research in network science to significantly advance the state-of-the-art in how networks influence and are influenced by human behavior in the context of military decision-making.